An automatic system for the detection of dairy cows lying behaviour in free-stall barns

Simona M.C. Porto, Claudia Arcidiacono, Umberto Anguzza, Andrea Giummarra, Giovanni Cascone

Department of Agri-food and Environmental Systems Management, Building and Land Engineering, University of Catania, Italy

Abstract

In this paper, a method for the automatic detection of dairy cow lying behaviour in free-stall barns is proposed. A computer vision-based system (CVBS) composed of a video-recording system and a cow lying behaviour detector based on the Viola Jones algorithm was developed. The CVBS performance was tested in a head-to-head free stall barn. Two classifiers were implemented in the software component of the CVBS to obtain the cow lying behaviour detector. The CVBS was validated by comparing its detection results with those generated from visual recognition. This comparison allowed the following accuracy indices to be calculated: the branching factor (BF), the miss factor (MF), the sensitivity, and the quality percentage (QP). The MF value of approximately 0.09 showed that the CVBS missed one cow every 11 well detected cows. Conversely, the BF value of approximately 0.08 indicated that one false positive was detected every 13 well detected cows. The high value of approximately 0.92 obtained for the sensitivity index and that obtained for QP of about 0.85 revealed the ability of the proposed system to detect cows lying in the stalls.

Introduction

Nowadays, the possible occurrence of pandemics caused by the diffusion of diseases transmitted from animals to humans represents one of the most threatening risk for human health. Therefore, operators and researchers of the livestock and veterinary sectors as well as the public opinion assign a great importance to those specific problems of intensive livestock farming which affect animal welfare, product quality and, as a consequence, consumers' health.

Recently, modern Information and Communication Technology (ICT) applications to animal breeding produced the development of the Precision Livestock Farming (PLF) which regards the study of systems and methods to monitor, store and continuously assess data related to processes regarding the animal and its breeding environment. The modelling of these data makes it possible to obtain information suitable to assess animal health status, comfort and productivity, and adopt smart management protocols of livestock farming in real time (Cangar *et al.*, 2008).

In dairy cow farming, PLF may represent a useful application for the study of animal behaviour. In this context, this study aims at designing and implementing a system, which was based on the automatic analysis of digital images, for behaviour detection of dairy cows housed in free-stall barns.

The automatic analysis of digital images has been utilised in several research work (Shao *et al.*, 1998; Noldus *et al.*, 2001; Shao and Xin, 2008) that had the objective to monitor animal behaviour in different breeding environments. Automatic systems of detection based on this technique require the adoption of image processing algorithms when relevant brightness and background variations reduce detection of the significant content of the image (foreground), i.e., the presence of the animals and the activities they carry out. Furthermore, the methodologies proposed in these studies regarded laboratory trials or tests in small breeding environments which entailed that the designed systems were very simple since the video-recording system included only one camera or video-camera, and the automatic detection system elaborated images acquired from only one source.

The main objective of this research, which was part of a wider project (PRIN 2008 - Innovative techniques for behaviour analysis of dairy cows in barns equipped with automatic milking systems), was the attainment of the automatic detection of some behaviours of dairy cows bred in a free-stall barn by using classifiers based on the Viola Jones algorithm (Viola and Jones, 2001, 2004). This algorithm, which was originally elaborated for human face detection, could be suitable to obtain an accurate recognition also in presence of relevant varia-

tions of brightness and background in the analysed image sequence.

The behaviours considered in this research were the following: feeding at the manger, lying in the stalls, and standing and walking in the feeding alley and the service alleys. For each behaviour, one or more classifiers were modelled.

In this paper the main information related to the modelling of the classifiers that were built to detect the presence of dairy cows lying within the stalls is reported. Furthermore, in order to improve the readability of the paper, a brief description of the Viola Jones algorithm is provided in the section below. Further information can be found in Porto *et al.* (2011, 2013).

Training and execution of a classifier based on the Viola Jones algorithm

The classifier based on the Viola Jones algorithm (Viola and Jones, 2001, 2004) is a *cascade* of stages, each one of them being a strong classifier composed of a combination of weak classifiers.

In the training phase, the classifier requires samples of positive images and samples of negative images that both have sizes of $w \times h$ pixels. To obtain these samples from the frames acquired by the video-cameras, it is required that an operator extracts a set of positive images that has a size of $W \times H$ pixels (with $W > w$ and $H > h$) and a set of negative images that has a size of $W' \times H'$ pixels (with $W' > w$ and $H' > h$). A positive image contains the object to be recognized whereas a negative image contains only parts of the objects that constitute the image background. In the extraction process, all the positive images must maintain the same aspect ratio w/h. The sizes of the negative images, on the other hand, must not be smaller than those of the largest positive image.

The sample of positive images is obtained by automatically reducing all the positive images selected by the operator to $w \times h$ pixels. The sample of negative images, instead, is generated by the algorithm through the automatic execution of multiple scans of the negative images selected by the operator. In detail, a first scan is carried out by using a sliding window of $w \times h$ pixels.

During the training process, the algorithm classifies a negative image as false positive if it has wrongly recognised the target object in the image. Likewise, the algorithm classifies a positive image as true positive if it has correctly recognised the target object in the image. Therefore, along with the maximum number of stages (*ns*) that compose the cascade, the training phase requires the definition of the maximum value for the false positive rate (*MaxFPR*), and the minimum value for the true positive rate (*MinTPR*). The *MaxFPR* and *MinTPR* fixed values must be equal for all stages.

The *MinTPR* parameter is obtained from the equation:

$$MinTPR = \left(TPR_{cas} \right)^{1/ns} \tag{1}$$

by fixing the minimum value of the true positive rate that the cascade must produce (*TPRcas*), and the value of *ns*.

At the end of the training process the minimum false positive rate of the cascade (*FPRcas*) is obtained. If *FPRcas* is of the order of magnitude 10^{-6} (Viola & Jones, 2001, 2004), it demonstrates that the classifier has the ability to distinguish the background from the object to be detected.

In the execution phase, the Viola Jones algorithm utilises an image scanning process by using a sliding window similar to that used in the training phase. Ultimately, each stage of the cascade establishes if the area of the image contained within the sliding window must be classified as *not an object* or as *probably an object*. If one of the stages establishes that the content is classified as *not an object*, that part of the

image is definitively discarded and classified as background, whereas if it is classified as *probably an object* the next stage will analyse the content. The higher the number of stages used to evaluate the content of the sliding window the higher the possibility that the sliding window contains the object to be detected.

Materials and methods

The breeding area under study

The research trials were conducted in a cubicle free-stall barn for dairy cows located in the province of Ragusa (Sicily, Italy). A group of 15 cows were monitored. The breeding environment considered in the trials was composed of a resting area subdivided in two rows of head-to-head stalls with sand beds, a feeding alley adjacent to the resting area, a service alley, and two service passages (Figure 1). In this breeding environment, the factors that may complicate the automatic detection of the lying cows within the stalls are the following: variability of colour tone of the sand which constitutes the lying surface of the stalls; the considerable variability of environment brightness in the proximity of the open side of the building; the high quantity of solar radiation that is reflected by the sand beds and the metal surfaces constituting the cross bars; the lack of homogeneity in colour and reflectance of the alleys' surfaces, which is caused by the presence of slurry.

To design and build the video-recording system, video-cameras Vivotek FD7131 were chosen, which have a maximum resolution of 640 × 480 pixels and a maximum image-capture capability of 30 fps, equipped with an hypertext transfer protocol interface (HTTP) and a LED (light emitting diode) for illumination in night vision. With the aim to obtain a planimetric top-view of the breeding environment considered in the trial, ten video-cameras were utilised, six of them were placed above the feeding alley at an height of 4.40 m from the floor, and four were installed above the resting area at an height of 4.05 m (Figure 1). Image acquisition was carried out by utilising a specific software developed in Microsoft® Visual C# Express, which allowed the acquisition of one snapshot for each video-camera with a frequency of one fps. Through the additional use of OPENCV graphical libraries, the software applied operations of synchronization, calibration, rotation, resizing and union to the ten obtained snapshots, in order to produce as output an image which had a size of 1044×1920 pixels and contained the plan view. The outcome of these operations was a sequence of panoramic top-view images with a frequency of 0.5 fps.
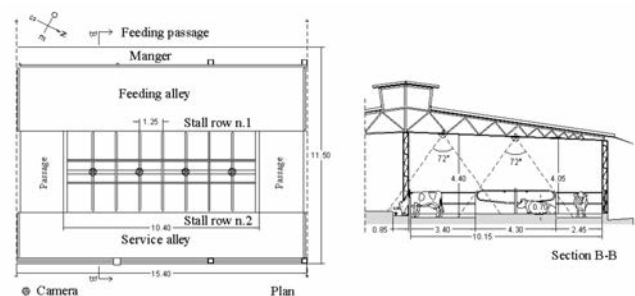

Figure 1. Plan and section of the study area in the barn considered in the trial.

Training of the classifiers

The panoramic images, which were utilized to train the cow lying behaviour detector, were extracted from the video-recordings acquired between the 1st and the 7th of August 2011.

Though the camera model was equipped with a built-in LED illuminator, it did not provide sufficient lighting inside the building during the evening and night. Therefore, the panoramic top-view images used in the training phase referred to the time interval between 6:00 a.m. and 8:00 p.m. However, this time interval did not affect the modelling of the detectors because it included the most significant herd management tasks and variations in the surfaces' illumination. In particular, it included feeding distribution, the two milking operations during which the animals left the resting area, the activation of the evaporative cooling devices and the direct dripping system that produced a significant variation in the colour of the sand beds in the stalls, the cleaning of the feeding alley that determined a variation of the appearance of the floor, and light intensity variations in the different areas of the breeding environment throughout the day.

Since the stalls had a head-to-head layout and a classifier based on Viola Jones algorithm is unable to find the object to be detected if its image is rotated compared to that used during the training, in this phase two classifiers were modelled: classifier 1a to detect the presence of cows lying in the first row of stalls which can be reached from the feeding alley, and classifier 1b to detect the presence of cows when they are lying in the second row of stalls which can be reached from the service alley.

By applying a sampling time of ten minutes, 84 panoramic top-view images of the resting area were considered in each of the seven days analysed. From these panoramic images an operator selected the positive and negative images required to train the classifiers.

Testing of the classifiers

The objective of the classifiers' testing phase was to verify the functioning of each trained classifier without performing time-consuming visual recognition activities required to validate the detection results.

As for the training phase, the panoramic images were extracted from video-recordings acquired between the 1st and the 7th of August 2011 from 6:00 a.m. to 8:00 p.m., with a 10-min scan sampling. However, acquisition was delayed by five minutes in comparison to that employed in the training phase. For each classifier and each instant of acquisition, the panoramic image was selected on the basis of the average pixel brightness level of the image of the stalls. In detail, the panoramic image that had the average of stalls' pixel brightness values closest to the weekly mean, which was obtained by averaging the brightness values of the same instant of acquisition for each of the seven days of the week, was selected.

As for the training phase, an operator extracted a set of positive images and a set of negative ones. The images utilised for the testing phase were built by a specific software that allowed the overlapping of each positive image on each one of the panoramic negative images, i.e., selected among those where the cows were absent from the resting area. The overlapping area was randomly selected among the areas of the negative images that corresponded to the stalls. The obtained test images were then subject to three image alteration techniques: smoothing, erosion, and dilatation. Finally, the software stored the coordinates of the vertices of each positive image by fixing one of the vertices of the panoramic negative image as origin of the coordinate system. Within the same software, two classifiers were implemented to assess the accuracy of the detection of the cows that were present in the test images. During the execution of the classifiers, the software labelled a test image as true positive (TP) if the cow was correctly detected in its position, or else as false negative (FN) if the cow though it was present in the test image, was not detected by the classifier. Furthermore, the software assigned a value equal to the number of false positives (FPs) eventually found, i.e., areas of the background which were wrongly detected as cows, to each test image. Next, the software counted the images labelled as TP and FN, computed the sum of the FPs found in the images, and quantified the following accuracy indices: i) Hit rate (HR): ratio between the total number of TPs and the total number of the images utilised in the testing phase; ii) Miss rate (MR): ratio between the total number of FNs and the total number of the images utilised in the testing phase; iii) False positive rate (FPR): ratio between the total number of FPs and the total number of the images utilised in the testing phase.

Validation of the system

Once verified the functioning of the two classifiers by means of the testing phase, the overall performance of the CVBS was assessed by finding the position of the cows displayed in the sequence of the panoramic images. To this aim, another software that allowed the comparison between the data obtained from the simultaneous execution of the two classifiers and those derived from the visual analysis of the panoramic images carried out by the operator was elaborated. This comparison made it possible to compute the following accuracy indices:

i) Branching factor (BF): ratio between the total number of FPs and the total number of TPs obtained from the panoramic images. Low values of the BF index indicate a good ability of the classifier to discern the cows from the background:

$$BF = \frac{FP}{TP} \qquad (2)$$

ii) Miss factor (MF): ratio between the total number of FNs and the total number of TPs obtained from the panoramic images. Low values of the index show the good ability of the classifier to recognize the cows:

$$MF = \frac{FN}{TP} \qquad (3)$$

iii) Sensitivity: percentage of cows correctly classified, (i.e., true positives), over the total number of cows which are actually present in the panoramic images:

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \qquad (4)$$

iv) Quality Percentage (QP): in comparison to the sensitivity index, this ratio takes into account the presence of FPs in the analysed panoramic image, i.e., areas of the background incorrectly classified as cows:

$$QP = \frac{TP}{TP + FN + FP} \times 100 \qquad (5)$$

The validation phase was carried out by using the video recordings acquired between the 8th and the 14th of August, 2011, during the time interval between 6:00 a.m. and 8:00 p.m. An operator performed the visual analysis of the video-recordings by using the software and marked the stalls occupied by the animals by selecting a point inside it. This information was then stored in a database. To reduce the time needed to perform the visual analysis, the panoramic images were selected at a 10-min scan sampling.

The number of positive images used in the training phase, equal to 826 and 319 for classifier 1a and classifier 2a respectively, was obtained by extracting from each panoramic image the image subsets of all the rectangular areas of 224×140 pixels, related to the stalls occupied by

OPEN ACCESS

the cows (Figure 2A). Similarly, the number of negative images was equal to 600 and was obtained from stalls areas unoccupied by the cows (Figure 2B). The dimensions of the sliding window were set to 40×25 pixels. In all, 600 positive images samples and about $50×10^7$ negative images samples of 40×25 pixels were obtained.

During the testing phase, the use of weekly data of stalls' pixel brightness values (Figure 3) produced a sequence of video-recordings composed of 84 panoramic images from which 346 positive images of 224×140 pixels were extracted. In detail, 274 images obtained by selecting all the cows lying in the stalls of row n.1 were utilised to test the classifier 1a. The other 72 images obtained from images where the cows were lying in the stalls of row n.2, were used to test the classifier 1b. As for the number of negative images, two images were selected at 6:30 a.m. and at 6:30 p.m., i.e., when the cows were in the milking area, other two images were selected slightly before and after the cleaning of the feeding alley, which was carried out at about 8:00 a.m., when the cows were confined in the service alley (Figure 1).

Results and discussion

For both the classifiers the *MaxFPR* value, which was set equal to 0.5, allowed a rapid decrease of the FPR_{cas} value as the number of stages increased, with a maximum number of stages *ns* equal to 30.

Under the hypothesis that the TPR_{cas} value was at least equal to 0.90, the parameter *MinTPR* computed according to Eq. (1), was equal to 0.9965. The training of the classifiers ended at the 27th stage when FPR_{cas} values reached the order of magnitude 10^{-7} for both classifiers.
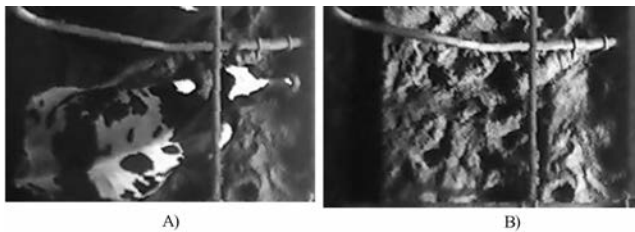

Figure 2. A) Example of positive image. B) Example of negative image
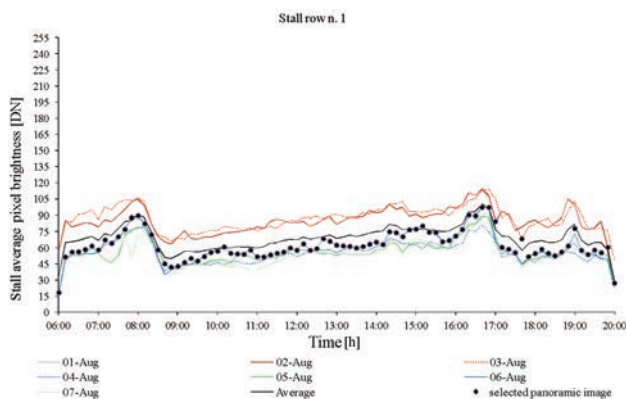

Figure 3. Average of the pixel brightness values composing the areas of the panoramic images of stall row n.1, computed at 10-min intervals and for each acquisition instant of the seven days considered; weekly means of the average of the pixel brightness values related to the stall row n.1; panoramic images selected to produce the sequence of images.

Similarly to the outcomes of a previous study carried out to detect cow presence at the feed barrier (Porto *et al.*, 2011), in this trial a lower number of both positive and negative images than that utilised by Viola and Jones (2001, 2004) for human face detection, made it possible to obtain very low FPR_{cas} values equal to $3.09×10^{-7}$ and $7.93×10^{-7}$ for classifier 1a and classifier 1b, respectively, and at the same time assured considerably high TPR_{cas} values equal to approximately 0.90 for both classifiers. This result could be explained by the fact that the classifiers modelled in this study, thanks to the use of Viola Jones algorithm, proved to be robust with regard to the significant background variations which occurred in the breeding environment during the daytime hours.

The operations of overlay and alteration of the images determined the use of 4384 test images for classifier 1a and 1152 for classifier 1b.

On the basis of the high value of HR which resulted about equal to 0.88 for both the classifiers, it can be highlighted that the testing phase proved the high ability of the classifiers to detect the cows even when noise affects the processed images, still yielding a low FPR which resulted equal to about 0.08 for both the classifiers. This result confirmed the outcome of the training phase where a very low value of FPR_{cas} was obtained.

In the validation phase, 589 panoramic top-view images were analysed in which 2281 images of cows lying in the cubicles were present. The overall number of cows correctly detected was 2088 and, therefore, the number of unclassified cows resulted 193. The number of false positives given by the system was 167. The accuracy indices calculated on the basis of these values were: BF = 0.08; MF = 0.09; sensitivity = 0.92; QP = 0.85.

The values obtained for the BF and MF indices showed that for every 11-13 cows correctly detected, the cow lying behaviour detectors made two errors: an omission error, i.e., they did not detect a cow in the resting area, and a commission error, i.e. they produced a false positive. However, the good ability of the classifiers to detect the cows is proved by the high value of the sensitivity obtained. This index which is directly comparable to the HR value obtained in the testing phase, revealed that in the validation phase both the classifiers improved of about 4% their ability to detect the cows in comparison to the testing phase. This improvement in the validation phase was not found in a previous trial carried out with the aim to detect the cows at the feed barrier (Porto *et al.*, 2011). This result is due to the smaller number of postures the animals assume while lying in the stalls compared to that found when they are in the feeding alley, close to the feed barrier. Furthermore, the stalls geometrically identify the area where cow presence is possible and thus it facilitates the selection of positive images which contain the body of just one animal.

By considering that the operations carried out to obtain the plan view of the resting area produced sequences of panoramic top-view images with a frequency of 0.5 fps and that the time needed to classify the content of a panoramic top-view image using an Intel Core (TM) 2 Quad CPU Q670 processor was approximately 320 ms, the automated detection of a cow in the stalls can be considered to be a real-time operation. During the execution of the cow lying behaviour detectors, this characteristic allowed results to be stored within text files, or in a database, therefore avoiding the storage of video-recordings.

Conclusions

This study aimed at building an automatic system for the detection of cows lying in the stalls. To this purpose two classifiers were modelled, which were based on the algorithm proposed by Viola and Jones. Since the proposed system did not require the use of image quality

enhancement techniques applied to the outcome of the classifiers, both in the training and the execution phases, its application could be more suitable than that of other systems proposed in literature when the images of the breeding environment do not allow an easy extraction of the animal shape from the background.

The results obtained in the training phase highlighted that the use of a small number of both positive and negative images makes it possible to obtain a low false positives rate, while maintaining a good cow detection ability within the stalls. The importance of this result lies in the reduction in the operator's time required to select positive and negative images used for the training of the classifiers.

The results of the testing phase proved that the two classifiers have a good ability to detect cows even when noise affects the processed images, and making a few errors in the background.

Finally, since the classification speed of the information contained in the panoramic images allowed the recognition of the dairy cows in real time, the system was able to store the results of the detection in text files or in databases that allow an easy data management, and thus avoids the onerous storage of video-recordings.

References

Cangar, Ö., Leroy, T., Guarino, M., Vranken, E., Fallon, R., Lenehan, J., Mee, J., Berckmans D. 2008. Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis. Computers and electronics in agriculture, 64, 53-60.

Noldus, L. P. J. J., Spink, A. J., Tegelenbosch, R. A. J. 2001. Ethovision: a versatile video tracking system for automation of behavioral experiments. Behavior Research Methods, Instruments, & Computers, 33(3), 398-414.

Porto, S.M.C., Arcidiacono, C., Guarnera, G. C., Cascone, G. 2011. Preliminary study for the implementation of an image analysis algorithm to detect dairy cow presence at the feed barrier. Journal of Agricultural Engineering, 4, 17-23.

Porto, S.M.C., Arcidiacono, C., Anguzza U., Cascone, G. 2013. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns. Biosystems Engineering, 115 (2), 184-194.

Shao, B., Xin, H., Harmon, J.D. 1998. Comparison of image feature extraction for classification of swine thermal comfort behaviour. Computers and Electronics in Agriculture, 19, 223-232.

Shao, B., Xin, H. 2008. A real-time computer vision assessment and control of thermal comfort for group-housed pigs. Computers and electronics in agriculture, 2008, 62, 15- 21.

Viola, P., Jones, M. 2001. Rapid object detection using a boosted cascade of simple features. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, 1, 511-518.

Viola, P., Jones, M. 2004. Robust real-time face detection. International Journal of Computer Vision, 57(2), 137-154